LANGLEY GRANT

7N-60-CR

114131

P-30

ACCURATE LOW-COST METHODS FOR PERFORMANCE EVALUATION OF CACHE MEMORY SYSTEMS

Subhasis Laha, Janak H. Patel and Ravishankar K. Iyer

Coordinated Science Laboratory
University of Illinois
1101 W. Springfield Ave.
Urbana, IL 61801

Abstract : Trace-driven simulation is a simple way of evaluating cache memory systems with varying hardware parameters. But to evaluate realistic workloads, simulating even a few millions of addresses is not adequate and such large scale simulation is impractical from the consideration of space and time requirement. In this paper, new methods of simulation based on joint usage of statistical techniques and the principles of computer architecture are proposed for decreasing the need for large trace measurements and to predict true program behavior. In our method, sampling techniques are applied while collecting the address trace from a workload; this drastically reduces the space needed to store the trace. New simulation techniques are developed to use the sampled data to predict not only the mean miss rate of the cache, but also its true distribution. Finally, a model is proposed to statistically project the results to different context-switch intervals from only one simulation of a small number of samples of a fixed size.

Acknowledgements : This research was supported in part by the Joint Services Electronics Program (US Army, US Navy, US Air-force) under contract N00014-84-C-0149 and in part by NASA grant NAG-1-613.

ACCURATE LOW-COST METHODS FOR PERFORMANCE EVALUATION OF CACHE MEMORY SYSTEMS

1. Introduction

It is well known that the use of a cache in memory hierarchy has considerable influence on the overall performance of a computer system. Although, cache memory systems have been extensively measured and simulated for traditional high level languages like Fortran, Pascal, Algol and PL1 [1], studies on LISP programs are few and far between [2, 3, 4].

The goal of this project was to develop new techniques for cache evaluation. For various reasons, discussed in the next section, trace driven simulation was selected as the method of evaluation. But, soon it was found that conventional simulation suffers from a lot of limitations; the results, so obtained, do not represent true program behavior. Some of these aspects have already been pointed out by Smith [3]. Since results on LISP program architecture are few, LISP programs were employed in our case studies. The proposed methodology, however, is equally applicable to other programs.

This paper establishes a new method of trace driven simulation for cache memory systems, based on sampling techniques. This method decreases the need for large trace measurements and simulation, and is based on joint usage of statistical techniques and the principles of computer architecture. It is shown that, unlike conventional techniques, the proposed approach is able to characterize a complete program with low overhead. Thus, it not only allows a prediction of the *mean miss rate*, but also of its *true distribution*. In the next section, different aspects of conventional trace driven simulation is presented along with its limitations.

2. Different Methods of Evaluation

Among the various methods of performance evaluation of cache based systems, *Hardware Monitoring* is the most accurate method. But it is very difficult to instrument and it takes up a lot of time to develop. Even after implementing it for a particular machine, it is hard to vary the parameters of the architecture. *Analytical Methods* can be quiet cheap in most cases. But it involves so many simplifying assumptions that the results may be far from being true. This leaves us with *Simulation* which is simple and quiet easy to develop. It can be used to evaluate any existing or proposed architectures. The parameters of the architecture can be varied very easily simply by changing the input parameters of a simulator.

2.1. Trace Driven Simulation

A program address trace consists of a sequence of the virtual (usually) addresses, referenced by a program or programs, along with other relevant information. Trace driven simulation is the process of running the simulation model of a system with the trace. A simulator can also be driven by a random number generator; but, because of the lack of existence of an accurate behavioral model of programs in the context of cache, such simulation can never represent realistic program behavior. On the other hand, a trace represents the behavior of at least one program and so, trace driven simulation will characterize that correctly. For all these reasons, this has been chosen by most people, including us for evaluation of memory systems.

2.1.1. Limitations

In spite of all these nice features, there are several limitations of trace driven simulation which make it inadequate not only for our purpose, but also as a tool of performance evaluation in general.

(1) It simulates only a very small part of a program. For example, if a program runs for a thousand seconds of cpu time at the rate of one million memory references per second, simulating a trace of one million addresses will only represent 0.01% of the whole program.

(2) Traces are usually collected from the initial portion of a program, rather than from the core part of it. This practice of taking a single section of a program and projecting it for the entire program is highly questionable. Smith [1, 3] has reported that an enormous degree of variability exists among traces in their measured miss ratios in cache. Our own experience with simulation of cache memories has shown that there is a large degree of variation in program behavior in one sample of half million references with another of the same program at a different time. Even if the mean miss rate, measured this way, happens to be the same as that of the whole program, this mean value is not adequate to accurately reflect the performance of the system. We will show that the distribution of this measured miss ratio is far from the real one. Such experience suggests that simulation with a section of program chosen on an ad-hoc basis with a few thousand references cannot accurately reflect the overall program behavior.

(3) Any large scale simulation with a few millions of addresses requires an enormous space to store the trace. It is also very expensive from the consideration of computer time. Each such simulation with one set of hardware parameters as input typically takes up several minutes of cpu time on a minicomputer (e.g., VAX-11/780). So, running a bunch of simulation jobs with several combinations of architectural parameters for a single trace is computationally very expensive.

(4) Most machines switch context every few thousands of instructions, especially in a multiprogramming environment. The Context switch can occur due to I/O interrupts or time out of a fixed time slice. This effect of context switch on cache performance has been studied by others [5, 6, 7]. But, their methods involve simulating every different switch intervals or a curve fitting on individual program traces. Our objective is to perform only one simulation, based on our sampling technique and statistically project the results for other switch intervals.

To overcome these limitations, we propose a new sampling-based technique for studying cache performance in LISP systems. The methodology is based on statistical methods and is capable of characterizing a program without extensive simulation.

3. Sampling Technique

In conventional simulation, the effect of task switches on a program can be estimated by assuming the cache to be flushed totally during a context switch [5, 8]. This is essentially true for smaller caches. For larger caches, the miss ratio obtained in this way provide a lower limit to the cache performance. The preliminary aim of our sampling technique is to provide a cost-effective approach to estimate the distribution of this cold-start miss ratio.

Instead of simulating all consecutive intervals of address references, as is done conventionally, we sample the address trace (Fig. 1) and run the simulation only on the sampled data.

Our Methodology

(1) Choose a sample size which is typically a few thousands of consecutive addresses. In the initial study, this corresponds to the task interval considered.

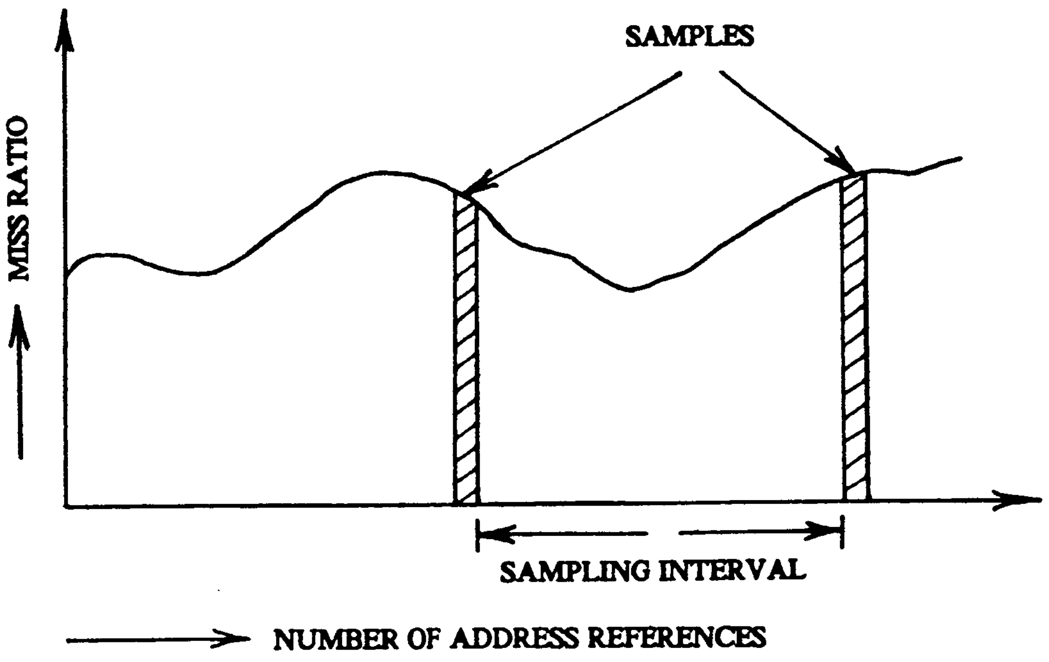

Fig. 1. Sampling technique

(2) Depending on the size of the whole trace, determine the uniform sampling interval to attain a certain number of samples.

(3) Sample the continuous trace accordingly (or, preferably collect only the sampled part of the trace at the beginning itself) and store the sampled addresses only.

(4) Assume cache to be empty at the beginning of each sample and simulate cache memory system for each sample.

Claim

The distribution of the average miss ratio for the samples gives an estimate of that of the cold-start miss ratio over the entire trace for the same context switch interval as the sample size.

Validation Procedure

Simulating one sample of consecutive addresses represents the behavior of that particular context switch interval. Because of the cold-start, the miss rate is very high at the beginning of the interval and usually decreases as the cache gets more and more filled up. We consider the average value of the miss ratio at the end of the sample. To validate this methodology, it is essential to establish that the miss ratios of the samples truly represent the behavior of the continuous trace. Our objective is to determine the number of samples which is adequate to represent any program, irrespective of the nature of the program or the length of the trace. We will also show that the number of samples is the factor to be considered for sampling, rather than the sampling interval. Therefore, to validate this technique, different number of samples are considered over different LISP programs and statistical techniques are applied to compare the distribution of the sampled data with that of the actual distribution of the miss ratio. In the next section, different aspects of the simulations performed are outlined.

4. Simulations Performed

4.1. Traces Used

The traces of virtual addresses were collected by running the compiled code of *Franz LISP programs* on a VAX-11/780. These are continuous traces which were later sampled at various intervals and sizes and then simulated. The programs chosen are realistic workloads of varied nature. The nature of the programs and the lengths of the corresponding traces are listed below:

(1) FSIM : Fault Simulation program : Length $\simeq$ 2.9 million.

(2) DRC : Design Rule Check program for VLSI layout : Length $\simeq$ 5.7 million.

(3) ELI : English Language Interpreter program : Length $\simeq$ 1.8 million.

(4) RRL : Mathematical Rewrite Rule Lab program : Length $\simeq$ 2.0 million.

These traces are collected by running only part of these programs because of practical limitations. But they are of varied length and we will show that our methodology holds good independent of this variation of trace length, so that this method can be applied to traces of entire programs also.

4.2. Size and Number of Samples

Three different sample sizes were used : 5K, 10K and 20K of addresses. They are chosen to be of the same order as the typical task interval on VAXes [9].

It is known that about 30 samples give a good estimate of any normally distributed data. At the beginning, we started with the assumption that the distribution of the cache miss ratio is normal which is found later not to hold good in general. But, nonetheless, the numbers of samples selected are (1) about 20, (2) about 35. The results of these two sampled cases are compared with that of the continuous trace. We will show that about 35 samples are always found adequate to predict the actual miss distribution with reasonable accuracy.

4.3. Architecture Simulated

The architectural model has three level memory hierarchy : virtual, main and cache. We consider *real address cache*: the virtual address is first translated to real address of the main memory which is subsequently mapped to the cache address space.

The main memory is organized into pages. The physical page frames are numbered sequentially as they are allotted. FIFO (First In First Out) policy is used for page replacement. The cache memory is *set-associative* and its blocks are replaced according to *Least Recently Used (LRU)* policy. Cache is updated by *no-allocate* policy i.e., during a write-miss, only the main memory is updated and it is not counted as a cache miss [10].

The goal of this paper is to develop a new method of simulation which is applicable to all LISP programs, rather than to evaluate the behavior of LISP programs for various architectures. For that reason, it is sufficient to simulate this architecture only with one set of hardware parameters. The sizes chosen for the three levels of memory, virtual, main and cache are respectively 64KB, 16KB and 4KB. The page size in the main memory is 128 bytes. The block size and set size of the cache are kept fixed at 8 bytes and 2.

4.3.1. Approximation

Both the instructions and the data in VAX-11 are of variable length. It has been shown in other studies [9] that most of the times, these lengths are within 4 bytes . Also, the VAX machine always fetches a block of 4 bytes together from the memory. So, it is quite reasonable to approximate the length of all the references to 4 bytes. This simplifies the process of gathering the trace and others [4, 11] have also made similar assumptions . During this simulation, if the address referenced fits exactly in an integral 4-byte boundary, only one block (of size at least 4 bytes) is fetched. Otherwise if the 4-byte field, starting from the current address in the trace, does not coincide with a 4-byte boundary and it also surpasses the end of the current block of the cache, then the next block is also fetched. In the next section, the results of simulation are presented and discussed.

5. Results

Recall that the goal is to show that a sampling scheme can successfully predict not only the true mean value but also the distribution of miss ratio. Note that it is assumed that a program is running with context switches. To establish the effectiveness of the sampling scheme, we compare the distribution of the sampled data with that of the continuous trace. As a first step, the mean and the standard deviation of the miss ratios of sampled data are compared with actual values (Table 1). The term "all" for the number of samples in the table signifies the results for the continuous trace without any sampling. It is found that the maximum error in mean miss ratio for number of samples $\simeq$ 35 is 5.5%; this error for about 20 samples can be as high as 13.8%. The standard deviation of miss ratio for about 35 samples is always very close to its actual value. So, just from the consideration of mean and standard deviation of the samples, it can be concluded that with approximately 35 samples a good estimate of the mean and the standard deviation is achieved.

The distribution of the continuous trace is compared with that obtained from sampled trace in Figures 2 through 5. All these figures contain the histograms for the case of about 35 samples and the continuous trace. Figures 2 and 3 also include the case for number of samples $\simeq$ 20. Of all the 12 cases considered, only four are given in these figures; remaining of them appear in the Appendix. It is very apparent from these figures that in general, the distribution profile is far from normal. Initially, these distributions can be just checked for similarity in appearance. The cumulative percentages of the sample points in the different levels, included in the figures, provide additional way of comparing the distribu-

Table 1. Mean & Standard Deviation of Sampled Result

Program	Sample size	Mean miss ratio (for #samples)			%Error in mean (for #samples)		Std. dev. of miss ratio (for #samples)		
		≃ 20	≃ 35	All	≃ 20	≃ 35	≃ 20	≃ 35	All
FSIM	5K	.0743	.0767	.0781	4.9	1.8	.0158	.0177	.0160
	10K	.0576	.0609	.0600	4.0	-1.5	.0137	.0115	.0131
	20K	.0492	.0513	.0505	2.6	-1.6	.0125	.0098	.0107
DRC	5K	.0768	.0639	.0675	-13.8	5.3	.0539	.0443	.0430
	10K	.0630	.0550	.0582	-8.2	5.5	.0425	.0386	.0389
	20K	.0576	.0512	.0528	-9.1	1.9	.0389	.0349	.0353
ELI	5K	.0819	.0775	.0792	-3.4	2.1	.0452	.0379	.0406
	10K	.0680	.0691	.0673	-1.0	-2.7	.0359	.0351	.0343
	20K	.0606	.0616	.0604	-0.3	-2.0	.0295	.0311	.0307
RRL	5K	.0424	.0417	.0416	-1.9	-0.2	.0354	.0332	.0329
	10K	.0349	.0356	.0346	-0.9	-2.9	.0283	.0293	.0286
	20K	.0318	.0315	.0308	-3.2	-2.3	.0267	.0258	.0255

(a) Number of samples = 20

```
MIDPOINT
   Y                                                        FREQ   CUM.    PERCENT     CUM.
                                                                   FREQ              PERCENT
0.028  |                                                       0      0       0.00      0.00
0.034  |*****                                                  1      1       5.00      5.00
0.040  |***********************************                    7      8      35.00     40.00
0.046  |********************                                   4     12      20.00     60.00
0.052  |***************                                        3     15      15.00     75.00
0.058  |                                                       0     15       0.00     75.00
0.064  |**********                                             2     17      10.00     85.00
0.070  |*****                                                  1     18       5.00     90.00
0.076  |**********                                             2     20      10.00    100.00
       -----+----+----+----+----+----+----+
            1    2    3    4    5    6    7
                   FREQUENCY
```

(b) Number of samples = 36

```
MIDPOINT
   Y                          FREQ   CUM.    PERCENT     CUM.
                                     FREQ              PERCENT
0.028  |                         0      0       0.00      0.00
0.034  |****                     2      2       5.56      5.56
0.040  |********                 4      6      11.11     16.67
0.046  |******************       9     15      25.00     41.67
0.052  |******************       9     24      25.00     66.67
0.058  |************             6     30      16.67     83.33
0.064  |******                   3     33       8.33     91.67
0.070  |**                       1     34       2.78     94.44
0.076  |****                     2     36       5.56    100.00
       ---+---+---+---+--
          2   4   6   8
          FREQUENCY
```

(c) Continuous Trace

```
MIDPOINT
   Y                                                       FREQ   CUM.    PERCENT     CUM.
                                                                  FREQ              PERCENT
0.028  |**                                                    2      2       1.40      1.40
0.034  |*******                                               7      9       4.90      6.29
0.040  |************************                             24     33      16.78     23.08
0.046  |**********************************                   34     67      23.78     46.85
0.052  |*****************************                        29     96      20.28     67.13
0.058  |*********************                                21    117      14.69     81.82
0.064  |****************                                     16    133      11.19     93.01
0.070  |*                                                     1    134       0.70     93.71
0.076  |*********                                             9    143       6.29    100.00
       -----+----+----+----+----+----+----
            5   10   15   20   25   30
                   FREQUENCY
```

Fig. 2. Frequency bar chart of miss ratio for FSIM : Context-switch interval = 20K

(a) Number of samples = 20

```
MIDPOINT
   Y                                                        FREQ   CUM.    PERCENT    CUM.
                                                                   FREQ              PERCENT
0.015  |***************                                       3      3      15.00      15.00
0.033  |******************************                        6      9      30.00      45.00
0.051  |                                                      0      9       0.00      45.00
0.069  |*****                                                 1     10       5.00      50.00
0.087  |**********                                            2     12      10.00      60.00
0.105  |                                                      0     12       0.00      60.00
0.123  |********************                                  4     16      20.00      80.00
0.141  |**********                                            2     18      10.00      90.00
0.159  |**********                                            2     20      10.00     100.00
       |
       -----+----+----+----+----+----+
            1    2    3    4    5    6
                 FREQUENCY
```

(b) Number of samples = 36

```
MIDPOINT
   Y                                                FREQ   CUM.    PERCENT    CUM.
                                                           FREQ              PERCENT
0.015  |************                                  6      6      16.67      16.67
0.033  |************************                     12     18      33.33      50.00
0.051  |**                                            1     19       2.78      52.78
0.069  |**                                            1     20       2.78      55.56
0.087  |**********                                    5     25      13.89      69.44
0.105  |********                                      4     29      11.11      80.56
0.123  |**********                                    5     34      13.89      94.44
0.141  |****                                          2     36       5.56     100.00
0.159  |                                              0     36       0.00     100.00
       |
       ----+---+---+---+---+---+
           2   4   6   8  10  12
               FREQUENCY
```

(c) Continuous Trace

```
MIDPOINT
   Y                                      FREQ      CUM.    PERCENT    CUM.
                                                    FREQ              PERCENT
0.015  |*********                          174       174     15.34      15.34
0.033  |*******************                384       558     33.86      49.21
0.051  |*                                   17       575      1.50      50.71
0.069  |*                                   25       600      2.20      52.91
0.087  |*******                            148       748     13.05      65.96
0.105  |*******                            148       896     13.05      79.01
0.123  |*******                            139      1035     12.26      91.27
0.141  |***                                 67      1102      5.91      97.18
0.159  |**                                  32      1134      2.82     100.00
       |
       -----+----+----+----
          100  200  300
          FREQUENCY
```

Fig. 3. Frequency bar chart of miss ratio for DRC : Context-switch interval = 5K

(a) Number of samples = 35

```
MIDPOINT
   Y                                                    FREQ   CUM.    PERCENT     CUM.
                                                               FREQ              PERCENT
 0.007     |************                                  3      3       8.57       8.57
 0.021     |****                                          1      4       2.86      11.43
 0.035     |****************************                  7     11      20.00      31.43
 0.049     |****                                          1     12       2.86      34.29
 0.063     |********                                      2     14       5.71      40.00
 0.077     |****************                              4     18      11.43      51.43
 0.091     |****************************                  7     25      20.00      71.43
 0.105     |********************************              8     33      22.86      94.29
 0.119     |********                                      2     35       5.71     100.00
           |
           ----+---+---+---+---+---+---+---+
               1   2   3   4   5   6   7   8
                     FREQUENCY
```

(b) Continuous Trace

```
MIDPOINT
   Y                                      FREQ   CUM.    PERCENT     CUM.
                                                 FREQ              PERCENT
 0.007     |*********                      17     17       9.71       9.71
 0.021     |****                            7     24       4.00      13.71
 0.035     |****************               32     56      18.29      32.00
 0.049     |***                             5     61       2.86      34.86
 0.063     |*****                           9     70       5.14      40.00
 0.077     |***************                30    100      17.14      57.14
 0.091     |*******************            38    138      21.71      78.86
 0.105     |***************                29    167      16.57      95.43
 0.119     |****                            8    175       4.57     100.00
           |
           -----+----+----+----
               10   20   30
                 FREQUENCY
```

Fig. 4. Frequency bar chart of miss ratio for ELI : Context-switch interval = 10K

(a) Number of samples = 33

```
MIDPOINT
   Y                                                     FREQ   CUM.    PERCENT     CUM.
                                                                FREQ              PERCENT
0.0145   |******************                                9      9      27.27      27.27
0.0295   |************************                         12     21      36.36      63.64
0.0445   |**********                                        5     26      15.15      78.79
0.0595   |****                                              2     28       6.06      84.85
0.0745   |                                                  0     28       0.00      84.85
0.0895   |****                                              2     30       6.06      90.91
0.1045   |**                                                1     31       3.03      93.94
0.1195   |**                                                1     32       3.03      96.97
0.1345   |**                                                1     33       3.03     100.00
         ----+---+---+---+---+---+
             2   4   6   8  10  12
               FREQUENCY
```

(b) Continuous Trace

```
MIDPOINT
   Y                                                     FREQ   CUM.    PERCENT     CUM.
                                                                FREQ              PERCENT
0.0145   |*********************                           107    107      27.23      27.23
0.0295   |**************************                      129    236      32.82      60.05
0.0445   |****************                                 80    316      20.36      80.41
0.0595   |*                                                 5    321       1.27      81.68
0.0745   |**                                               10    331       2.54      84.22
0.0895   |**                                               12    343       3.05      87.28
0.1045   |******                                           31    374       7.89      95.17
0.1195   |***                                              13    387       3.31      98.47
0.1345   |*                                                 6    393       1.53     100.00
         ----+---+---+---+---+---+--
            20  40  60  80 100 120
               FREQUENCY
```

Fig. 5. Frequency bar chart of miss ratio for RRL : Context-switch interval = 5K

tions. It is found that the distributions of nearly 20 samples have significant difference from the true ones. But, the profiles of about 35 samples closely resemble the actual shape in all different cases. It is easily shown that the distributions of the samples are statistically equivalent to the actual ones with a high degree of confidence. Therefore, it is concluded that 35 samples are adequate for this method.

The distribution of the miss ratios provides a lot of information about the program behavior which is lost when only the mean value is considered. For example, the distribution of miss ratio is useful in performance analysis of cache memories in multiprocessors or cache memories with a finite write-back buffer. In both cases, a high miss rate imposes a relatively higher penalty *per miss* than the penalty per miss at low miss rates. These distributions for the four programs are quite different from one another; but, for a particular program, it follows similar trend for the various context switch intervals (i.e., for different sample sizes in our method). The miss ratio distribution of the DRC program has two distinct modes while for the FSIM program, it is nearest to the normal distribution among the four programs. This indicates that the miss ratio of the DRC program is spread over two regions which are apart from each other; the region with smaller value corresponds to the tight loop structures in the program. The miss ratio versus time plot (Fig. 6) for this program corroborates this feature. The FSIM program on the other hand has its miss ratio randomly spread over a single band as found in its miss ratio versus time plot.

5.1. Comparison with Conventional Simulation

In conventional simulation, the measured distribution of the miss ratio can be quite different from the actual one. The conventional method for cold-start miss ratio can be considered as simulating consecutive samples with no interval in between. To compare conventional simulation with our method, the distribution of miss ratio of 35 consecutive context switch intervals of size 5K addresses is checked for closeness with the actual

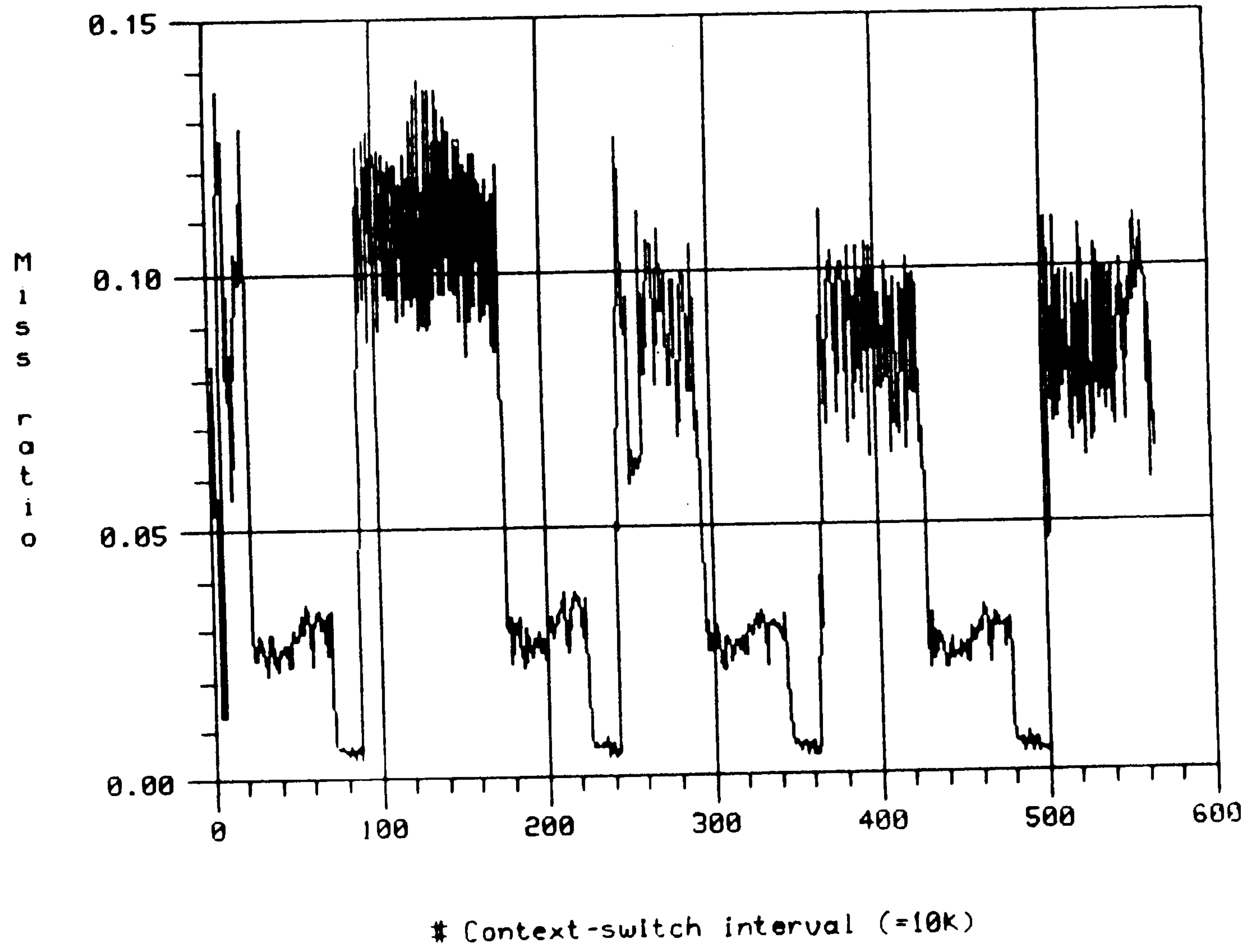

Fig. 6. Miss ratio versus time : DRC program : Context switch interval = 10K

distribution of miss ratio for DRC program, because this simulation has the same space-time cost as our method. These two distributions are found to differ considerably in mean value, standard deviation and the shape (Fig. 7). This is expected from Figure 6 which shows that 35 consecutive intervals, each of size 5K cover only a small, initial part of the program which is very different from the rest. This shows the inadequacy of conventional method to represent true program behavior.

6. Evaluation of Other Context Switch Intervals

In this section, we propose a model of cache behavior for larger context switch intervals, based on the performance measures of smaller intervals. Our aim is to perform only one simulation on samples of one size and statistically project the results for other context switch intervals. Once simulation is performed for samples of a particular size, information about smaller task intervals is already present in the results. So, only those intervals which are greater than the sample size need to be projected. To do that, we will first consider how the mean value of miss ratio over the samples vary within the duration of the sample.

Figure 8 shows the typical plot of mean miss ratio (over different samples) with the number of address references during the period of the sample size. The value of mean miss ratio at each point is calculated by considering the number of hits and misses from the end of the interval, rather than from the beginning for each sample. This is done to avoid the effect of cold start at the beginning of the interval on the later segments of the task interval. This enables us to check if the instantaneous mean miss rate in an interval reaches a steady value as the cache gets more and more filled up. It is observed in all our experiments that the miss ratio really attains a steady state for sufficiently long samples (as small as 5K

(a) Conventional method with 35 intervals

(Mean : 0.083; Standard deviation : 0.048)

```
MIDPOINT
   Y                                                       FREQ   CUM.   PERCENT    CUM.
                                                                  FREQ             PERCENT
 0.015  |*************************                            5      5    14.29     14.29
 0.033  |*****                                                1      6     2.86     17.14
 0.051  |******************************                       6     12    17.14     34.29
 0.069  |***********************************                  7     19    20.00     54.29
 0.087  |***************                                      3     22     8.57     62.86
 0.105  |***************                                      3     25     8.57     71.43
 0.123  |*****                                                1     26     2.86     74.29
 0.141  |********************                                 4     30    11.43     85.71
 0.159  |*************************                            5     35    14.29    100.00
        |
        -----+----+----+----+----+----+----+
             1    2    3    4    5    6    7
                     FREQUENCY
```

(b) Complete Trace

(Mean : 0.068; Standard deviation : 0.043)

```
MIDPOINT
   Y                          FREQ    CUM.   PERCENT    CUM.
                                      FREQ             PERCENT
 0.015  |*********             174     174    15.34     15.34
 0.033  |*******************   384     558    33.86     49.21
 0.051  |*                      17     575     1.50     50.71
 0.069  |*                      25     600     2.20     52.91
 0.087  |*******               148     748    13.05     65.96
 0.105  |*******               148     896    13.05     79.01
 0.123  |*******               139    1035    12.26     91.27
 0.141  |***                    67    1102     5.91     97.18
 0.159  |**                     32    1134     2.82    100.00
        |
        -----+----+----+----
            100  200  300
             FREQUENCY
```

Fig. 7. Comparison of conventional method for DRC : Context-switch interval = 5K

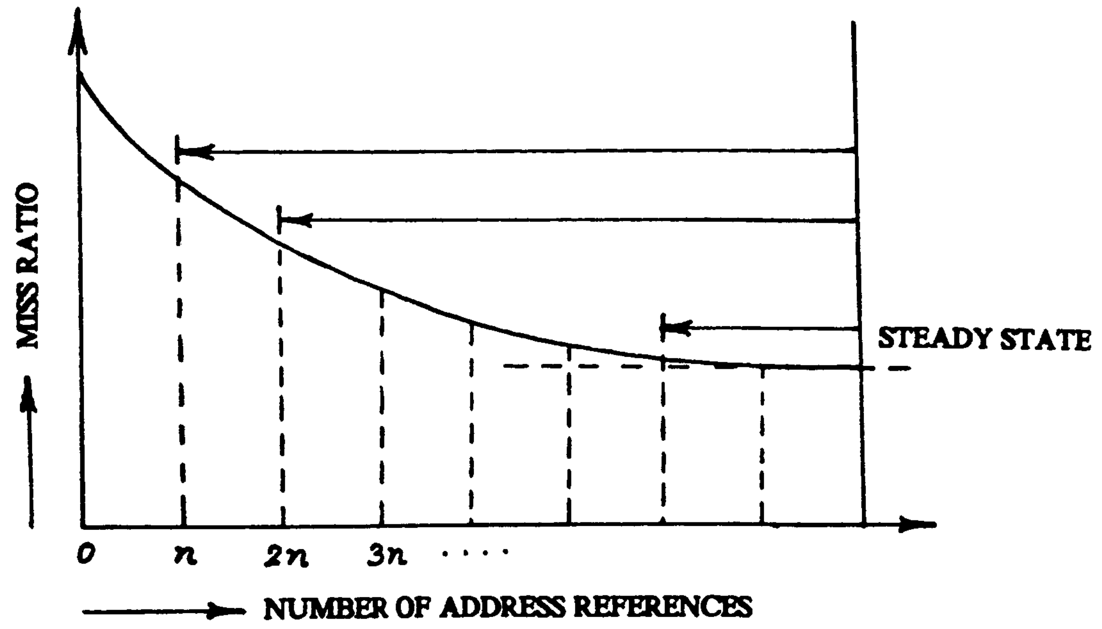

Fig. 8. Mean miss ratio versus number of addresses

for all the traces).

6.1. Our Model

Given a sample size n_0 which is long enough for the miss ratio to reach a stable value, this steady state will continue for the rest of a larger task interval n (Fig. 9).

Based on this assumption, the value of the miss ratio for a larger context switch interval n can be easily calculated and is given by :

$$m(n) = \frac{n_0}{n} m(n_0) + (1 - \frac{n_0}{n}) m_{ss}$$

where, m(n) : miss ratio for interval n

$m(n_0)$: miss ratio for interval n_0

m_{ss} : steady state value of miss ratio

6.2. Results

Based on the formula above, the miss ratio for 20K interval is predicted from that of 5K and 10K samples for all the traces; also, the miss ratio for 10K context switch interval is calculated from that of 5K samples (Table 2). It is found that the projected value is reasonably close to the actual value, measured from the continuous trace. However, we are still in the process of comparing the distributions of the projected miss ratio with the actual ones. More exhaustive statistical analysis needs to be done to find out how much can the samples be extended to larger context switch intervals with reasonable accuracy and to resolve other related issues.

7. Concluding Remarks

In this paper, we have established a sampling-based technique for cache memory simulation. This method provides insight into the true program behavior by giving the

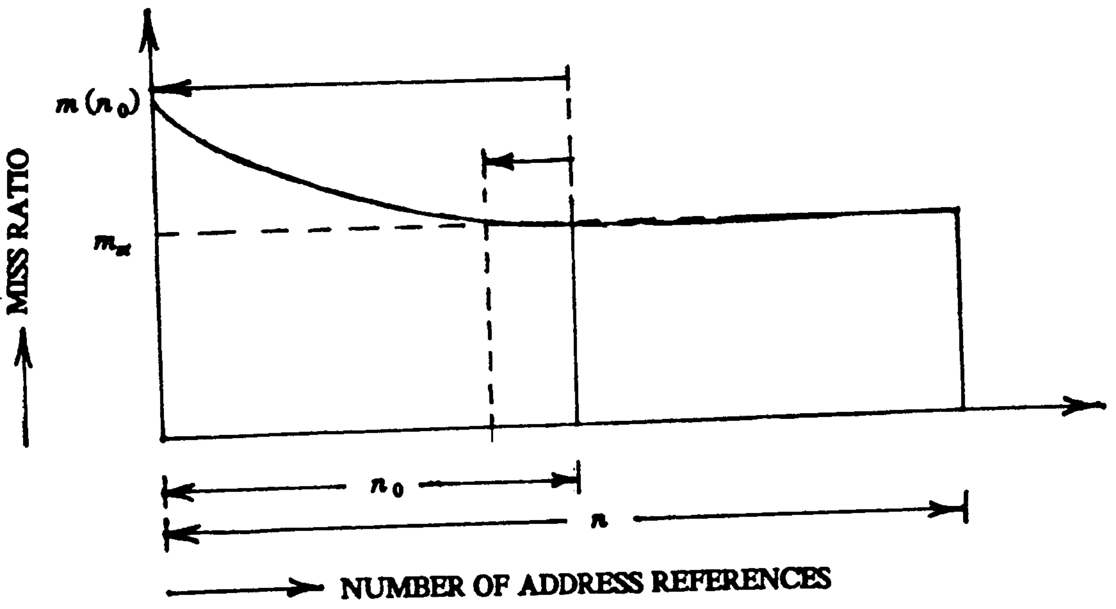

Fig. 9. Model for larger context switch intervals

Table 2. Calculated Miss Ratios for Different Context Switch Intervals

Program	Context switch interval	Actual measured miss ratio	Miss ratio calculated from context switch interval	
			5K	10K
FSIM	10K	.060	.062	
	20K	.051	.053	.054
DRC	10K	.057	.061	
	20K	.052	.056	.0525
ELI	10K	.067	.067	
	20K	.061	.0615	.0635
RRL	10K	.034	.035	
	20K	.030	.0315	.032

accurate estimates of both the mean value of miss ratio and its distribution. In comparison, conventional techniques just measure the mean miss ratio for one small segment of a program; thus, they may not be representative of actual performance. The cost of the sampling-based method in terms of both memory space and computer time is small. There is one pathological case in which this technique might not give satisfactory result: this will happen very rarely when (1) the miss ratio versus time plot of the program has a step-function, as in Figure 6, (2) the period of this variation is constant and (3) the sampling interval happens to be the same as this period and all the samples are taken predominantly at the top or bottom level of the step-function. This can be avoided by taking the samples at random intervals or by using two sets of samples having two different uniform sampling intervals. Hence the paper shows that the sampling-based simulation is an accurate approach. More measurements with other programs are essential to establish guidelines for general usage.

Simulating only one set of samples and projecting the result for other context switch intervals look very promising. But, still some more analysis has to be done regarding this issue. Another direction in which this research needs to be extended is to separate the cold-start effect at the beginning of a task interval from the characteristics of the rest of the interval. Finally, these new simulation methods can be very effectively used not only for cache memory simulation of LISP programs, they can also be easily extended for other programming languages.

Acknowledgements

We would like to thank Hoichi Cheong, Susheel Chandra, Ram Kunda, Suresh Rajamoney and G. Sivakumar for their assistance in collecting the traces. Many thanks to

Ram Chillarege and Sandy Hsueh for helping in the statistical analysis.

References

[1] A. J. Smith, "Cache Memories," *ACM Computing Surveys*, vol. 14, 3, pp. 473-530, September, 1982.

[2] D. W. Clark, "Measurements of Dynamic List Structure Use in Lisp," *IEEE Trans. Software Eng.*, vol. SE-5, 1, pp. 51-59, January, 1979.

[3] A. J. Smith, "Cache Evaluation and the Impact of Workload Choice," *Proc. 12th Annual Int'l Symp. on Computer Architecture*, vol. 13, 3, pp. 64-73, June, 1985.

[4] A. J. Smith, "Line (Block) Size Choice for CPU Cache Memories," UC Berkeley CS Report, UCB/CSD85/239, June, 1985..

[5] W. D. Strecker, "Cache Memories for PDP-11 Family Computers," *Proc. Third Annual Conf. on Computer Architecture*, pp. 155-158, 1976.

[6] W. D. Strecker, "Transient Behavior of Cache Memories," *ACM Trans. Computer Systems*, vol. 1, 4, pp. 281-293, November, 1983.

[7] I. J. Haikala, "Cache Hit Ratios with Geometric Task Switch Intervals," *Proc. 11th Annual Int'l Symp. on Computer Architecture*, vol. 12, 3, pp. 364-371, June, 1984.

[8] M. Easton, "Computation of Cold Start Miss Ratios," *IEEE Trans. Comp.*, vol. C-27, 5, pp. 404-408, May, 1978.

[9] J. S. Emer and D. W. Clark, "A Characterization of Processor Performance in the VAX-11/780," *Proc. 11th Annual Int'l Symp. on Computer Architecture*, vol. 12, 3, pp. 301-310, June, 1984.

[10] K. Hwang and F. A. Briggs, *Computer Architecture and Parallel Processing*. McGraw-Hill, 1984.

[11] M. D. Hill and A. J. Smith, "Experimental Evaluation of On-Chip Microprocessor Cache Memories," *Proc. 11th Annual Int'l Symp. on Computer Architecture*, vol. 12, 3, pp. 158-166, June, 1984.

Appendix

(Frequency bar charts for remaining context switch intervals)

(a) Number of samples = 36

```
MIDPOINT
   Y                                                      FREQ  CUM.   PERCENT    CUM.
                                                                FREQ             PERCENT
 0.045   |                                                   0     0     0.00      0.00
 0.055   |************                                       6     6    16.67     16.67
 0.065   |************                                       6    12    16.67     33.33
 0.075   |********************                              10    22    27.78     61.11
 0.085   |**************                                     7    29    19.44     80.56
 0.095   |******                                             3    32     8.33     88.89
 0.105   |****                                               2    34     5.56     94.44
 0.115   |**                                                 1    35     2.78     97.22
 0.125   |**                                                 1    36     2.78    100.00
         ----+----+----+----+----+
             2    4    6    8   10
               FREQUENCY
```

(b) Continuous Trace

```
MIDPOINT
   Y                                                      FREQ  CUM.   PERCENT    CUM.
                                                                FREQ             PERCENT
 0.045   |**                                                12    12     2.09      2.09
 0.055   |***********                                       57    69     9.93     12.02
 0.065   |*************************                        125   194    21.78     33.80
 0.075   |************************                         118   312    20.56     54.36
 0.085   |***********************                          115   427    20.03     74.39
 0.095   |******************                                88   515    15.33     89.72
 0.105   |*********                                         45   560     7.84     97.56
 0.115   |**                                                12   572     2.09     99.65
 0.125   |                                                   2   574     0.35    100.00
         ----+----+----+----+----+----+-
            20   40   60   80  100  120
                   FREQUENCY
```

Fig. 10. Frequency bar chart of miss ratio for FSIM : Context-switch interval = 5K

(a) Number of samples = 36

```
MIDPOINT
   Y                                                      FREQ  CUM.   PERCENT    CUM.
                                                                FREQ             PERCENT
 0.0335  |                                                   0     0     0.00      0.00
 0.0405  |******                                             3     3     8.33      8.33
 0.0475  |******                                             3     6     8.33     16.67
 0.0545  |**************                                     7    13    19.44     36.11
 0.0615  |************************                          12    25    33.33     69.44
 0.0685  |**********                                         5    30    13.89     83.33
 0.0755  |****                                               2    32     5.56     88.89
 0.0825  |******                                             3    35     8.33     97.22
 0.0895  |**                                                 1    36     2.78    100.00
         ----+----+----+----+----+----+
             2    4    6    8   10   12
                  FREQUENCY
```

(b) Continuous Trace

```
MIDPOINT
   Y                                                      FREQ  CUM.   PERCENT    CUM.
                                                                FREQ             PERCENT
 0.0335  |****                                               8     8     2.79      2.79
 0.0405  |**************                                    27    35     9.41     12.20
 0.0475  |******************                                36    71    12.54     24.74
 0.0545  |*****************************                     57   128    19.86     44.60
 0.0615  |*******************************                   62   190    21.60     66.20
 0.0685  |********************                              39   229    13.59     79.79
 0.0755  |******************                                35   264    12.20     91.99
 0.0825  |********                                          15   279     5.23     97.21
 0.0895  |****                                               8   287     2.79    100.00
         -----+-----+-----+-----+-----+-----+-
             10    20    30    40    50    60
                    FREQUENCY
```

Fig. 11. Frequency bar chart of miss ratio for FSIM : Context-switch interval = 10K

(a) Number of samples = 36

```
MIDPOINT
   Y                                                  FREQ  CUM.   PERCENT    CUM.
                                                            FREQ            PERCENT
 0.008  |**********                                      5     5    14.71     14.71
 0.024  |************************                       12    17    35.29     50.00
 0.040  |                                                0    17     0.00     50.00
 0.056  |**                                              1    18     2.94     52.94
 0.072  |******                                          3    21     8.82     61.76
 0.088  |************                                    6    27    17.65     79.41
 0.104  |******                                          3    30     8.82     88.24
 0.120  |****                                            2    32     5.88     94.12
 0.136  |****                                            2    34     5.88    100.00
        |
        ----+----+----+----+----+----+
            2    4    6    8   10   12
                 FREQUENCY
```

(b) Continuous Trace

```
MIDPOINT
   Y                                                  FREQ  CUM.   PERCENT    CUM.
                                                            FREQ            PERCENT
 0.008  |**************                                 35    35    13.01     13.01
 0.024  |****************************                   71   106    26.39     39.41
 0.040  |************                                   30   136    11.15     50.56
 0.056  |****                                           11   147     4.09     54.65
 0.072  |****                                            9   156     3.35     57.99
 0.088  |**************                                 35   191    13.01     71.00
 0.104  |***********                                    28   219    10.41     81.41
 0.120  |***************                                38   257    14.13     95.54
 0.136  |*****                                          12   269     4.46    100.00
        |
        ----+----+----+----+----+----+----+
           10   20   30   40   50   60   70
                 FREQUENCY
```

Fig. 12. Frequency bar chart of miss ratio for DRC: Context-switch interval = 10K

(a) Number of samples = 36

```
MIDPOINT
   Y                                                  FREQ  CUM.   PERCENT    CUM.
                                                            FREQ            PERCENT
 0.007  |********                                        4     4    11.76     11.76
 0.021  |************************                       12    16    35.29     47.06
 0.035  |**                                              1    17     2.94     50.00
 0.049  |**                                              1    18     2.94     52.94
 0.063  |******                                          3    21     8.82     61.76
 0.077  |****                                            2    23     5.88     67.65
 0.091  |************                                    6    29    17.65     85.29
 0.105  |**********                                      5    34    14.71    100.00
 0.119  |                                                0    34     0.00    100.00
        |
        ----+----+----+----+----+----+
            2    4    6    8   10   12
                 FREQUENCY
```

(b) Continuous Trace

```
MIDPOINT
   Y                                                  FREQ  CUM.   PERCENT    CUM.
                                                            FREQ            PERCENT
 0.007  |******************                             18    18    13.43     13.43
 0.021  |********************************               32    50    23.88     37.31
 0.035  |****************                               16    66    11.94     49.25
 0.049  |********                                        8    74     5.97     55.22
 0.063  |*                                               1    75     0.75     55.97
 0.077  |*********                                       9    84     6.72     62.69
 0.091  |********************                           20   104    14.93     77.61
 0.105  |**************************                     26   130    19.40     97.01
 0.119  |****                                            4   134     2.99    100.00
        |
        -----+----+----+----+----+----+--
             5   10   15   20   25   30
                 FREQUENCY
```

Fig. 13. Frequency bar chart of miss ratio for DRC: Context-switch interval = 20K

(a) Number of samples = 35

```
 MIDPOINT
    Y                                                          FREQ  CUM.   PERCENT    CUM.
                                                                     FREQ            PERCENT
0.0155  |****************                                        4     4     11.43     11.43
0.0325  |****************************                            7    11     20.00     31.43
0.0495  |****                                                    1    12      2.86     34.29
0.0665  |                                                        0    12      0.00     34.29
0.0835  |************************                                6    18     17.14     51.43
0.1005  |********************************                        8    26     22.86     74.29
0.1175  |****************************                            7    33     20.00     94.29
0.1345  |********                                                2    35      5.71    100.00
0.1515  |                                                        0    35      0.00    100.00
         ----+---+---+---+---+---+---+---+
             1   2   3   4   5   6   7   8
                 FREQUENCY
```

(b) Continuous Trace

```
 MIDPOINT
    Y                                                          FREQ  CUM.   PERCENT    CUM.
                                                                     FREQ            PERCENT
0.0155  |***************                                        37    37     10.57     10.57
0.0325  |****************************                           69   106     19.71     30.29
0.0495  |******                                                 15   121      4.29     34.57
0.0665  |******                                                 15   136      4.29     38.86
0.0835  |**************                                         36   172     10.29     49.14
0.1005  |*******************************                        77   249     22.00     71.14
0.1175  |***************************                            68   317     19.43     90.57
0.1345  |**********                                             26   343      7.43     98.00
0.1515  |***                                                     7   350      2.00    100.00
         ----+----+----+----+----+----+----+---
            10   20   30   40   50   60   70
                 FREQUENCY
```

Fig. 14. Frequency bar chart of miss ratio for ELI : Context-switch interval = 5K

(a) Number of samples = 35

```
 MIDPOINT
    Y                                     FREQ  CUM.   PERCENT    CUM.
                                                FREQ            PERCENT
0.006  |******                              3     3      8.57      8.57
0.018  |                                    0     3      0.00      8.57
0.030  |****************                    8    11     22.86     31.43
0.042  |**                                  1    12      2.86     34.29
0.054  |**                                  1    13      2.86     37.14
0.066  |****                                2    15      5.71     42.86
0.078  |********************               10    25     28.57     71.43
0.090  |****************                    8    33     22.86     94.29
0.102  |****                                2    35      5.71    100.00
        ---+---+---+---+---+
           2   4   6   8  10
             FREQUENCY
```

(b) Continuous Trace

```
 MIDPOINT
    Y                                           FREQ  CUM.   PERCENT    CUM.
                                                      FREQ            PERCENT
0.006  |********                                  8     8      9.20      9.20
0.018  |**                                        2    10      2.30     11.49
0.030  |****************                         17    27     19.54     31.03
0.042  |****                                      4    31      4.60     35.63
0.054  |*                                         1    32      1.15     36.78
0.066  |*******                                   7    39      8.05     44.83
0.078  |****************************             28    67     32.18     77.01
0.090  |**************                           14    81     16.09     93.10
0.102  |******                                    6    87      6.90    100.00
        -----+-----+-----+-----+-----+---
             5    10    15    20    25
                 FREQUENCY
```

Fig. 15. Frequency bar chart of miss ratio for ELI : Context-switch interval = 20K

(a) Number of samples = 33

```
MIDPOINT
   Y                                                      FREQ  CUM.   PERCENT    CUM.
                                                                FREQ            PERCENT
0.007    |******************                                 9     9     27.27     27.27
0.021    |**                                                 1    10      3.03     30.30
0.035    |********************************                  16    26     48.48     78.79
0.049    |                                                   0    26      0.00     78.79
0.063    |**                                                 1    27      3.03     81.82
0.077    |**                                                 1    28      3.03     84.85
0.091    |********                                           4    32     12.12     96.97
0.105    |**                                                 1    33      3.03    100.00
0.119    |                                                   0    33      0.00    100.00
         ----+---+---+---+---+---+---+---+
             2   4   6   8  10  12  14  16
                 FREQUENCY
```

(b) Continuous Trace

```
MIDPOINT
   Y                                    FREQ  CUM.   PERCENT    CUM.
                                              FREQ            PERCENT
0.007    |***********                     52    52     26.53     26.53
0.021    |**                              12    64      6.12     32.65
0.035    |*******************             94   158     47.96     80.61
0.049    |                                 1   159      0.51     81.12
0.063    |                                 2   161      1.02     82.14
0.077    |**                               9   170      4.59     86.73
0.091    |***                             16   186      8.16     94.90
0.105    |**                               9   195      4.59     99.49
0.119    |                                 1   196      0.51    100.00
         ----+---+---+---+---
            20  40  60  80
             FREQUENCY
```

Fig. 16. Frequency bar chart of miss ratio for RRL : Context-switch interval = 10K

(a) Number of samples = 33

```
MIDPOINT
   Y                                                      FREQ  CUM.   PERCENT    CUM.
                                                                FREQ            PERCENT
0.008    |******************                                 9     9     27.27     27.27
0.018    |                                                   0     9      0.00     27.27
0.028    |********************************                  16    25     48.48     75.76
0.038    |**                                                 1    26      3.03     78.79
0.048    |                                                   0    26      0.00     78.79
0.058    |**                                                 1    27      3.03     81.82
0.068    |****                                               2    29      6.06     87.88
0.078    |****                                               2    31      6.06     93.94
0.088    |****                                               2    33      6.06    100.00
         ----+---+---+---+---+---+---+---+
             2   4   6   8  10  12  14  16
                 FREQUENCY
```

(b) Continuous Trace

```
MIDPOINT
   Y                                        FREQ  CUM.   PERCENT    CUM.
                                                  FREQ            PERCENT
0.008    |**************                     26    26     26.53     26.53
0.018    |*                                   1    27      1.02     27.55
0.028    |***************************        49    76     50.00     77.55
0.038    |**                                  3    79      3.06     80.61
0.048    |**                                  3    82      3.06     83.67
0.058    |*                                   1    83      1.02     84.69
0.068    |**                                  3    86      3.06     87.76
0.078    |***                                 5    91      5.10     92.86
0.088    |****                                7    98      7.14    100.00
         -----+-----+-----+-----+-----+
             10    20    30    40    50
               FREQUENCY
```

Fig. 17. Frequency bar chart of miss ratio for RRL : Context-switch interval = 20K